COO-3077-163

Courant Mathematics and Computing Laboratory

U.S. Department of Energy

ETHICAL PROCEDURES IN COMPARATIVE EVALUATION OF DRUGS

O. E. Percus
and
J. K. Percus

COO 3077 - 163

U.S. Department of Energy Report

Prepared under Contract DE-AC02-76ER03077
with the Office of Energy Research

Mathematics and Computing
February 1980

New York University

UNCLASSIFIED

Courant Mathematics and Computing Laboratory
New York University

Mathematics and Computing COO-3077-163

ETHICAL PROCEDURES IN COMPARATIVE EVALUATION OF DRUGS

O. E. Percus
and
J. K. Percus

February 1980

U. S. Department of Energy
Contract DE-AC02-76ER03077

UNCLASSIFIED

Ethical Procedures in Comparative Evaluation of Drugs

O.E. Percus

and

J.K. Percus[*]

1. Introduction

The comparison of sampled distributions is one of the classical problems of statistics. Evaluation of any technique for doing so depends at least implicitly on an assumed cost function. In one of the most important applications, that of choosing treatments for the ailments of man, costs are measured in suffering and death, and must be accorded primary importance. In this lecture, I would like to address a highly simplified but in fact prototypical situation of this kind. It is that in which two drugs, A and B, are available for a specified illness, their actions are unequivocal or dichotomic: success or failure, and this action is on a time scale short compared to that on which patients receiving the drugs present themselves. The assumption is that the mode of treatment with each drug is sufficiently **stylized that** each drug can be fully represented for each patient independently by success probability p. The problem then is to design a sequential selection of treatments so that the drug with the higher success probability is used as often as possible [1].

2. Basic Sampling

The classical sampling method is direct and basically unsophisticated. A large set of patients is treated simultaneously or sequentially - it does not matter - by random choice of one of the two drugs and the data is collected, resulting in $f_A n_A$ successes on n_A trials with drug A, $f_B n_B$ on n_B trials with drug B. If the unknown underlying success probabilities for the two drugs are p_A and p_B, the difference of the observed relative frequencies f_A and f_B has a mean and variance

$$\begin{aligned} E\,(f_A - f_B) &= p_A - p_B \\ \sigma^2(f_A - f_B) &= \frac{p_A q_A}{n_A} + \frac{p_B q_B}{n_B} \end{aligned} \tag{1}$$

where

$$q_A = 1 - p_A\ , \quad q_B = 1 - p_B$$

determining a normal distribution for large enough sample size:

$$P\ (f_A - f_B) = (2\pi\sigma^2)^{-1/2}\ \exp - (f_A - f_B - E)^2 / 2\sigma^2 \tag{2}$$

In particular,

$$P\ (f_A - f_B \geq f + E) = \frac{1}{\sqrt{2\pi}} \int_{\frac{f}{\sigma}}^{\infty} e^{-\frac{x^2}{2}}\, dx \tag{3}$$

Thus if the claim is made that $p_B \geq p_A$ or $E \leq 0$, and observation yields $f_A - f_B \geq f$, the claim is to be rejected, and hence $p_A > p_B$ accepted with a confidence

$$\rho = 1 - \frac{1}{\sqrt{2\pi}} \int_{\frac{f}{\sigma}}^{\infty} e^{-\frac{x^2}{2}}\, dx \tag{4}$$

For this purpose, of course, σ^2 is estimated as

$$\sigma^2 (f_A - f_B) \sim \frac{f_A(1 - f_A)}{n_A} + \frac{f_B(1 - f_B)}{n_B} \tag{5}$$

What does the knowledge that $p_A > p_B$ with high confidence mean? It presumably means that if we now use drug A for a large population of size M, the expected number of patients following the sampling who receive the wrong drug because of our misplaced confidence is $\sim M(1 - \rho)$, and hence the expected excess number of failures after sampling $C_{AS} \sim M(1 - \rho)(p_A - p_B)$. Evaluating (4) asymptotically for large n_A, n_B, or small σ, then

$$C_{AS} \sim \frac{M(p_A - p_B)f}{\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2}\frac{f^2}{\sigma 2}} \tag{6}$$

$$\sim k\,M\,n^{\frac{1}{2}} e^{-\alpha n}$$

for sample size $n_A + n_B = n$ and suitable constants k, α. To hold down this added cost C_{AS}, we would want to choose the sample size so that $e^{\alpha n} \sim M$, or

$$n \propto \ln M \tag{7}$$

We will next formalize this procedure.

3. Effect of Patient Horizon

We turn from the somewhat vague notion of confidence to the explicit one of cost, expressed as mean excess failure rate in treating a population, due to our initial lack of knowledge of the success probabilities of the drugs involved. To design an effective sampling strategy, we will assume some foreknowledge, however, of the total population N to be treated, the so-called "patient horizon". In the context of the classical sampling techniques, the problem then is to decide how extensive a period of random sampling is to be used before settling on the better drug and using it.

For simplicity, but not necessity, we assume that in the sampling period, drugs A and B have each been tried $\frac{n}{2}$ times, a decision made, and the drug of choice used on the remaining population of $M = N-n$ patients. Suppose that $p_A > p_B$ and that there are A successes with drug A, B with drug B during the sampling period. Then since B will be chosen if $B > A$ at the end of the sampling, the number of excess failures for the complete set of N patients is clearly given by the contribution C_S during sampling plus C_{AS} after sampling:

$$C(A,B) = p_A n - (A + B) + (p_A - p_B)(N - n)\,\theta\,(B - A)$$

$$\text{where}\quad \theta\,(X) = \begin{cases} 1 & \text{for}\quad X > 0 \\ 1/2 & \phantom{\text{for}}\quad X = 0 \\ 0 & \phantom{\text{for}}\quad X < 0 \end{cases} \tag{8}$$

Averaging over the binomial distribution for A and B, the mean excess failures are then

$$C = (p_A - p_B)\frac{n}{2} + \tag{9}$$

$$+ (p_A - p_B)(N-n) \sum_{A,B} \binom{n/2}{A}\binom{n/2}{B} p_A^A\, q_A^{n/2-A} \quad p_B^B\, q_B^{n/2-B}\ \theta(B-A)$$

To put C into a usable form, we may first express $\theta(B - A)$ as a contour integral

$$\theta(B - A) = \frac{1}{2\pi i}\oint e^{iS(B-A)}\,\frac{dS}{S} \tag{10}$$

after which the summations over A and B are easily performed, yielding

$$C = (p_A - p_B)\,\frac{n}{2} + \tag{11}$$

$$+ (p_A - p_B)(N-n)\frac{1}{2\pi i}\oint \frac{dS}{S}\left(p_A e^{-iS} + q_A\right)^{\frac{n}{2}}\left(p_B e^{iS} + q_B\right)^{\frac{n}{2}} \quad .$$

Unless n is small, a saddle point expansion in S is appropriate. The saddle point $\bar{S}$ is found to be

$$e^{i\bar{S}} = \delta^{-1/2} \qquad (12)$$

where $\delta = p_B q_A \,/\, p_A q_B$,

and so on substituting,

$$C \sim (p_A - p_B)\left(\frac{n}{2} + k(N-n)\,\gamma^n\right)$$

where $\gamma = \sqrt{p_A p_B} + \sqrt{q_A q_B}$ (13)

$$1/k = (p_A p_B q_A q_B \Pi^2)^{1/4} \ln\left(\frac{p_A q_B}{q_A p_B}\right) .$$

If N is fixed, the cost C can now be minimized by choice of the sample size n, and for N >> n we readily find

$$\text{Min } C \sim \frac{(p_A - p_B)}{2} \frac{\ln N}{\ln 1/\gamma} \qquad (14)$$

$$\text{at } \gamma^n \sim 1/(2Nk\ln{}^1/_\gamma) ,$$

consistent with the estimate given by (7). For a better feeling as to how this procedure is affected by the underlying probabilities p_A and p_B , suppose that $p_A - p_B = \Delta$ is small. Expanding in Δ , we have

$\gamma \sim 1 - \frac{\Delta^2}{8pq}$, $k \sim \Delta^{-1}\sqrt{pq/\pi}$, and so

if

$$p_A - p_B = \Delta \ll 1$$

$$\text{Min } C \sim 4pq \quad \frac{\ln N}{\Delta} \tag{15}$$

$$\text{at} \quad n \sim \frac{8pq}{\Delta^2} \ln N :$$

for small Δ , we of course need a longer sampling period, but more importantly, we are guaranteed to raise the total mean excess failures.

4. Controlled Sampling

We have thus far described a quite rigid strategy: given an estimate of the patient horizon (and an implied estimate of the range of success probabilities to be encountered) we fix an optimal sample size n and carry out the sampling accordingly. Thus, no advantage is taken of the particulars of the information actually obtained, which could very well allow a decision to be made at an earlier stage. This leads to the notion of a stopping rule, in which the sampling is terminated when sufficient information is available to make a decision according to some predetermined criterion. The criterion is based traditionally upon establishing a superior drug to at least some designated probability.

Two sampling strategies have received principal attention [2]: vector-at-a-time and play-the-winner. The first of these, vector-at-a-time or VT, is a slight modification of the random sampling we have discussed above. In VT, one creates equal samples of A and B by the fixed procedure of alternating drugs A and B : ABABAB... , so that the unit test is the pair AB. Again denoting the total successes with A by A and with B by B, a cutoff s is chosen, together with the

Stopping rule: Terminate the sampling
and choose drug A the first time that (16)
$A - B = s$ before $A - B = -s$,

and of course vice versa. The questions asked are then: What is the probability that a correct choice has been made? How long does one expect the sampling to take?

To answer these questions, we observe first that as the result of each pair of tests, A will have increased by 1 with probability p_A , B by 1 with probability p_B . Hence, simplifying notation by $p_A = p$, $1 - p_A = q$, $p_B = p'$, $1 - p_B = q'$, the change in $A - B$ is given by

$$\Delta(A-B) = \begin{cases} 1 & \\ 0 \text{ with probability} & pp' + qq' \\ -1 & p'q \end{cases} \quad pq' \text{ for } 1 \tag{17}$$

Furthermore, the probability of deciding on A as the drug of choice is

$$P\text{ (choose A)} = \sum_{n=1}^{\infty} P(A-B = s \text{ at the } n^{th} \text{ pair}, \; |A-B| \neq s \text{ at prior pairs}) \tag{18}$$

Suppose that the system is described by a state space of vectors $\{v(j)\}$ whose j^{th} component is the probability that $A - B = j$, and with t^{th} unit vector designated by u_t. Define $Q_{\pm}$ to project onto the complement of $A - B = \pm s$:

$$(Q_{\pm} v)(j) = (1 - \delta_{j,\pm s}) v(j) \quad . \tag{19}$$

Then if P is the transition probability matrix defined by (17)

$$P(j',j) = \begin{cases} pq' & \text{for} \quad j'-j = 1 \\ pp' + qq' & \qquad j'-j = 0 \\ p'q & \qquad j'-j = -1 \end{cases} \tag{20}$$

the first n-1 steps of (18) are given by Q_+Q_-P, and (18) transcribes at once to

$$\begin{aligned} P\ (\text{choose } A) &= \sum_n u_s' \, P(Q_+Q_-P)^{n-1} \, u_0 \\ &= u_s' \, P(I-Q_+Q_-P)^{-1} \, u_0 \end{aligned} \tag{21}$$

To evaluate (21), we write it as

$$P(\text{choose } A) = u_s' Pv$$

where (22)

$$v - Q_+Q_-Pv = us \ ,$$

or in component form, using $1 = pp' + qq' + pq' + p'q$,

$$\begin{aligned} |t| < s:\ & pq'\bigl(v(t) - v(t-1)\bigr) + p'q\bigl(v(t) - v(t+1)\bigr) = \delta_{t,0} \\ & v(s) = v(-s) = 0 \end{aligned} \tag{23}$$

The linear difference equation (23) has the characteristic polynomial $pq'(x - x^2) + p'q\ (x - 1) = 0$ and hence the elementary solutions

$$v(t) = x^{-t} \quad , \quad x = \begin{cases} 1 \\ \delta = p'q/q'p \end{cases} \tag{24}$$

Substituting into (23) we find

$$s \geq t \geq 0 : \ v(t) = \frac{1 - \delta^{st}}{pq'(1-\delta)(1+\delta^s)} \tag{25}$$

and hence conclude that

$$P(\text{choose } A) = (Pv)(s) = \frac{1}{1 + \delta^s} \tag{26}$$

i.e.

$$P(\text{correct choice}) \geq \frac{1}{1 + \varepsilon} \text{ if } \quad s \geq \frac{\ln 1/\varepsilon}{\ln 1/\delta} \tag{27}$$

The expected number of trials required to come to a decision is dirivable in similar fashion:

$$E(n) = \sum_n n u_s' P(Q_+Q_-P)^{n-1} u_0 \; / \; \sum_n u_s' P(Q_+Q_-P)^{n-1} u_0$$

$$= u_s'(I - Q_+Q_- P)^{-2} u_0 \; / \; u_s' P(I - Q_+Q_-P)^{-1} u_0 \tag{28}$$

$$= 1 + \frac{\partial}{\partial\gamma} \ln \left(u_s' P(I - \gamma Q_+Q_-P)^{-1} u_0\right)\Big|_{\gamma=1}$$

which works out to

$$E(n) = s \, \frac{1 - \delta^s}{1 + \delta^s} \quad . \tag{29}$$

5. Ethical Sampling

The stopping rule strategy is a first step in using the accumulated information during a sample to modify the nature of the sample. But it operates on a global basis: the sampled population still effectively constitutes a set of guinea pigs, although care is taken that no more guinea pigs be used than necessary. An ethical procedure for sequential trials would be to give each patient the benefit of the prior information, so that none serve merely as guinea pigs.

A step in designing such an ethical strategy is to at least have the drug administered to a given patient according to what happened to the preceding patient. This is the "play-the-winner" or PW scheme, in which a patient is given drug A either if A was tried and it succeeded or if B was tried and it failed; B is given in the reciprocal case. If the stopping rule (16) is now applied, the analysis proceeds as in the VT case, with some modification occasioned by the lack of a-priori knowledge of the drug to be given to the n^{th} patient. Thus the state is now built from unit vectors in which A-B is fixed and a specified drug is to be used next.

Hence we write

$$P_n(t,\alpha) = P\ (\text{arrive at } A-B=t \text{ after } n \text{ trials and will next use } \{{A \atop B} \text{ as } \alpha = \{{1 \atop 0}\ ). \tag{30}$$

The transition matrix is clearly given by

$$\begin{aligned} P_n(t,1) &= p\ P_{n-1}(t-1,1) + q'P_{n-1}(t,0) \\ P_n(t,0) &= q\ P_{n-1}(t,1) + p'\,P_{n-1}(t+1,0) \end{aligned} \tag{31}$$

for $n \geq 1$.

If the first patient will receive A or B with probability 1/2 then $P_0(t,\alpha) = 1/2\ \delta_{t,0}$ and equation (23) is now replaced by:

$$\begin{aligned} |t| < s : &\begin{pmatrix} v(t,1) \\ v(t,0) \end{pmatrix} - \begin{pmatrix} p & 0 \\ 0 & 0 \end{pmatrix}\begin{pmatrix} v(t-1,1) \\ v(t-1,0) \end{pmatrix} - \begin{pmatrix} 0 & q' \\ q & 0 \end{pmatrix}\begin{pmatrix} v(t,1) \\ v(t,0) \end{pmatrix} - \\ &\begin{pmatrix} 0 & 0 \\ 0 & p' \end{pmatrix}\begin{pmatrix} v(t+1,1) \\ v(t+1,0) \end{pmatrix} = \begin{pmatrix} 1/2 \\ 1/2 \end{pmatrix} \delta_{t,0} \\ &\begin{pmatrix} v(s,1) \\ v(s,0) \end{pmatrix} = 0 = \begin{pmatrix} v(-s,1) \\ v(-s,0) \end{pmatrix} \quad . \end{aligned} \tag{32}$$

Eliminating

$$v(t,0) = \frac{1}{q'} \left(v(t,1) - pv(t-1,1) - \frac{1}{2}\delta_{t,0}\right) \quad , \ |t|<s \tag{33}$$

and again using $pp' + pq' + p'q + qq' = 1$, while setting $v(t,1) \equiv v(t)$ we have

$$-s < t < s-1 : p(v(t) - v(t-1)) + p'(v(t) - v(t+1)) = \frac{1}{2}(1+q')\delta_{t,0} - \frac{1}{2}p'\delta_{t,-1} \tag{34}$$

$$v(-s) = 0 = v(s-1)$$

On solving (34) and substituting, there results

$$P(\text{choose } A) = (Pv)(s,1) = \frac{1 - (1+\mu)\lambda^s}{1 - \mu\lambda^s} \tag{35}$$

$$\lambda = {}^{p'}/_{p} \qquad \mu = {}^{q'}/_{q}$$

With VT sampling, the sample size directly determined the mean excess failures during sampling. That is, each pair of tests results in one use of the poorer drug and hence p-p' mean excess failures. From (29), the mean "cost" during sampling is then

$$C_{VT} = s(p - p') \frac{1 - \delta^s}{1 + \delta^s} \quad , \ \delta = {}^{p'q}/_{pq'} \tag{36}$$

To find the cost in the PW scheme, we can insert a matrix $\begin{pmatrix} 1 & 0 \\ 0 & \gamma \end{pmatrix}$ at each step to pick up a muliplicative factor γ each time drug B is tried. Hence we have to compute

$$C_{PW} = (p-p') \frac{\delta}{\delta\gamma} \ln \left[(1,0)\, P \left(I - Q_+ Q_- P \begin{pmatrix} 1 & 0 \\ 0 & \gamma \end{pmatrix} \right)^{-1} \begin{pmatrix} 1/2 \\ 1/2 \end{pmatrix} \right]_{\gamma=1} , \quad (37)$$

which works out to

$$C_{PW} = s \frac{(q + \frac{p}{2s})(1 - \lambda^s)(1 - p\lambda - q\lambda^s)}{1 - p\lambda - q\lambda^{2s}} , \quad (38)$$

$$\lambda = p'/p \quad .$$

6. Comparison

The leading question is whether an attempt to improve the chances of an individual patient, as in play-the-winner, results in a substantial disadvantage to the whole population. The answer to this question depends upon the point of view adopted. Suppose first that we want to distinguish between the two drugs with high probability, $P(\text{correct choice}) \geq P^*$, whenever the difference in effectiveness exceeds some thereshhold $p - p' \geq \Delta^*$. If the mean effectiveness

$$\bar{p} = \frac{1}{2}(p + p') \quad (39)$$

is fixed, this sets the stopping rule termination parameter s. From (26) ,

$$s_{VT} = \ln \frac{P^*}{1-P^*} \Bigg/ \ln \frac{(\bar{p} + \frac{1}{2}\Delta^*)(\bar{q} + \frac{1}{2}\Delta^*}{(\bar{p} - \frac{1}{2}\Delta^*)(\bar{q} - \frac{1}{2}\Delta^*)} \qquad (40)$$

and from (35)

$$s_{PW} = \ln \frac{2\bar{q} - (\bar{q} + \frac{1}{2}\Delta^*)P^*}{(\bar{q} - \frac{1}{2}\Delta^*)(1-P^*)} \Bigg/ \ln \frac{\bar{p} + \frac{1}{2}\Delta^*}{\bar{p} - \frac{1}{2}\Delta^*} \qquad (41)$$

Now, using (36) and (38), the ratio C_{PW} / C_{VT} can be plotted as a function of $\bar{p}$. Doing so, one finds [2] a relative advantage depending very much on where in $(\Delta^*, \bar{p})$ space the comparison is made. In particular, for $P^* \sim 1$ and $\Delta^* << 1$, PW is superior (has lower mean excess failures) only for $\bar{p} > \frac{3}{4} - \frac{1}{8}\Delta^*$, while VT is superior in the large low success rate region.

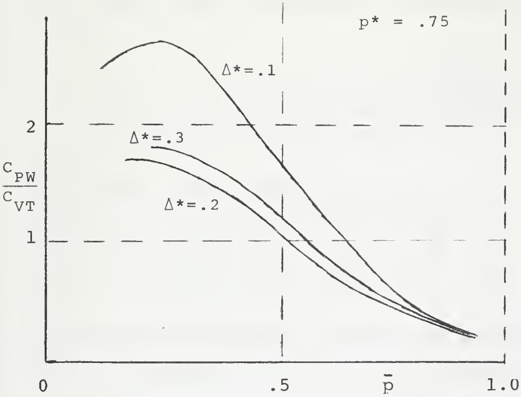

Figure 1.

Of course, as we have previously mentioned, the significance of mean excess failures during sampling (at fixed probability of choosing the better drug) is somewhat questionable, the mean total excess failure depending as it does on the total population treated. Thus, for estimated horizon N one is really interested in choosing the sample truncation parameter s to minimize

$$C = C(\text{sample}) + (N-\bar{n}(\text{sample})(1 - P(\text{choose } A))(p-p') \qquad (42)$$

where $\bar{n}$ is the mean sample size and A is assumed the better drug. If the pair (p,p') is actually known, but not its allocation to the two drugs, this results in a very different picture [3]

Table 1

Total Mean Excess Failure at $p - p' = 0.2$

Horizon	$\bar{p} = 0.2$			$\bar{p} = 0.5$			$\bar{p} = 0.8$		
	FS	VT	PW	FS	VT	PW	FS	VT	PW
10^2	14	2.8	2.6	23	3.8	3.4	14	2.8	2.0
10^4	28	6.6	5.8	46	10.4	8.0	28	6.6	4.0
10^6	41	10.0	8.8	68	16.0	13.2	41	10.0	6.0

In the range covered, PW is uniformly superior to VT, although not by much. Both, however, are much better than the fixed sample FS technique (14).

However, if only an upper bound on N is known, and a range of possible values of p and p', the choice of optimal s - to minimize the maximum C consistent with these possibilities-leads to a still different pattern as shown [3]

Table 2

Total Mean Excess Failure at $p - p' = 0.2$

s chosen to accord with $N \leq 10^6$, $0.3 \leq p \leq 0.9$

actual N	$\bar{p} = 0.2$		$\bar{p} = 0.5$		$\bar{p} = 0.8$	
	VT	PW	VT	PW	VT	PW
10^2	15	36	15	21	15	5.6
10^4	15	36	15	21	15	5.6
10^6	15	36	16	21	15	6

Under these circumstances, the "ethical" PW strategy is superior only for sufficiently high $\bar{p}$.

7. Improved Versions

Many modifications of the basic PW and VT schemes have been proposed, primarily with the objective of reducing $E(N_B)$, the expected number of trials with the inferior drug during the testing period. Perhaps the simplest, due to Hoel, Sobel & Weiss, [4] involves a "pretesting" period in which each drug is tested n_0 times, resulting in success numbers A_0, B_0 . If $\max(A_0,B_0) > k_c$, some critical value, PW is thereafter employed, otherwise VT. A quite small value of n_0 , say 5, is empirecally effective in reducing $E(N_B)$

at given (Δ^*, p^*).

Other modifications maintain PW strategy, but alter the stopping rule so that sampling also ends when each drug has exceeded a prescribed number of failures - Berry & Sobel [5] - or when the total failures exceed some cutoff - Fuschini [6].

A still more sophisticated stopping rule - Simon, Weiss & Hoel [7] - employs the likelihood function

$$L(p,p') = p^{s} p'^{s'} q^{f} q'^{f'} \tag{43}$$

for s, s' successes, f,f' failures at underlying probabilities p,p', and stops when the ratio

$$\Lambda = \mathcal{L}_1 / \mathcal{L}_2 \tag{44}$$

reaches a predetermined upperbound (or its reciprocal), where

$$\mathcal{L}_1(\Delta^*) = \max_{\Delta^* \leq p \leq 1} L(p, p-\Delta^*)$$

$$\mathcal{L}_2(\Delta^*) = \max_{\Delta^* \leq p \leq 1} L(p-\Delta^*, p)$$

In addition to using this in the context of PW, these authors have also suggested an "adaptive assignment" (with a stopping rule): try drug A with probability $\theta = \frac{\Lambda}{1+\Lambda}$, drug B with probability $1 - \theta = \frac{1}{1+\Lambda}$. Not only is a much improved strategy for low enough $\bar{p}$ achieved, but each patient during sampling also has the benefit of the prior history of patients.

One aspect of stopping rule strategies is worth noting. Since the probability of a correct choice on stopping after n trials is usually of the form $1 - \alpha^n$, and the number of poor drug choices during the trials proportional to n, it follows that if n is adjusted optimally for patient horizon N, both mean n and mean cost will go as lnN. This is a standard against which purported ethical strategies must be measured, a matter to which we now turn our attention.

8. Testing Without Termination

We have thus far laid great emphasis on the importance of terminating the sampling with despatch and deciding upon the better drug. With modern methods of information storage and retrieval, this philosophy is less compelling, and there is little reason not to go on "forever", while making the treatment strategy as ethical as possible. What strategies then are suggested? The PW scheme gives us an idea of what to expect as a result of unterminated sampling. Suppose to start that $p_A = 1$ but $p_B = p$ is arbitrary. Each time B is tried, we have a probable excess loss of $1 - p = q$; if A is tried, B is never tried again, leading to no increase of excess loss. Thus the only possibilites are that B is tried at the first step, with a probability of $\frac{1}{2}$, it succeeds s times for a probability of p^s , and then fails with probability q. Hence the asymptotic mean excess loss is $\frac{1}{2}(q + pq + p^2q + \ldots)q$, or

$$p_A = 1 \quad : \quad C_\infty = \frac{1}{2} q \tag{46}$$

far better thean the generic N^{th} stage mean excess loss $C_N \propto \ln N$.

On the other hand, suppose that $p_B < p_A < 1$. Now A has a non-zero probability of failing, so that B will keep on being tried. Only the limiting mean excess loss per trial over N trials will converge as $N \to \infty$

$$c_\infty = \lim_{N\to\infty} \frac{C_N}{N} = (p_A - p_B)P(\text{choose B}) \tag{47}$$

A run of exactly n trials of B occurs with probability $p_B^{n-1} q_B$ yielding a mean length of

$$1q_B + 2p_B q_B + 3p_B^2 q_B + \dots = 1/q_B \quad .$$

Hence, $P(\text{try } B) = \frac{1}{q_B} \Big/ \frac{1}{q_A} + \frac{1}{q_B} = q_A \Big/ q_A + q_B$, and we conclude that

$$c_\infty = \frac{q_B - q_A}{q_B + q_A} \, q_A \tag{48}$$

far worse asymptotically than the generic $c_N \propto \ln N/N$. In fact (48) has a maximum for $q_B > q_A$ of $c_\infty = 3 - 2\sqrt{2} = 0.172$ at $q_B = 1$, $q_A = \sqrt{2} - 1 = 0.414\dots$.

This attempt to introduce ethical considerations by at least retaining the result of the preceding treatment, if not terminated, hence exacts a large penalty (48) from the full population of patients. We now turn to the major thrust of this lecture: how does the situation change when each patient is given the advantage of the information accumulated by the result of testing all previous patients? Everything hinges of course on just how this information is used to advantage, and for this purpose we adopt a Bayesian point of view [8]. We suppose a symmetric a-priori distribution of the two success probabilities: $f_0(p_A,p_B)$. Then after $N = A + B + a + b$ trials with A,B successes, a,b failures, the a-posteriori $f_N(p_A,p_B|A,B,a,b)$ is determined by

$$\begin{aligned} f_N(p_A,p_B|A,B,a,b)f(A,B,a,b) &= f(p_A,p_B;A,B,a,b) \\ &= f(A,B,a,b|p_A,p_B)f_0(p_A,p_B) = \\ &= \binom{A+B}{A}\binom{a+b}{a} p_A^A \, p_B^B \, q_A^a \, q_B^b \; f_0(p_A,p_B) \end{aligned} \tag{49}$$

Hence

$$f_N(p_A,p_B|A,B,a,b) = n_{ABab} \, p_A^A \, p_B^B \, q_A^a \, q_B^b \, f_0(p_A,p_B) \tag{50}$$

for a suitable normalization constant n_{ABab} .

If the a-priori $f_0(p_A,p_B)$ is taken as a Beta-distribution.

$$f_0 \; (p_A,p_B) = n_0 \, p_A^e \, q_A^f \, p_B^e \, q_B^f \tag{51}$$

then this form is retained in f_N which is a computational advantage. Now we can write

$$f_{N}, (p_A,p_B|A,B,a,b) = n_N \, p_A^{A+e} \; q_A^{a+f} \; p_B^{B+e} \; q_B^{b+f} \tag{52}$$

This is as if there had been $2e + 2f$ trials before the first patient, with e successes and f failures for both drugs A and B. The advantage to be taken of (52) is clear enough. It follows at once from (52) that after N trials

$$\frac{E(p_A)}{E(q_A)} = \frac{A+e+1}{a+f+1} \qquad (53)$$

and similarly for drug B. We therefore adopt the strategy for each patient entering the sequence

$$\left.\begin{array}{c}\text{choose A}\\ \text{flip for A or B}\\ \text{choose B}\end{array}\right\} \text{ if } (A+e+1)(b+f+1)-(B+e+1)(a+f+1) \left\}\begin{array}{l}>0\\ =0\\ <0\end{array}\right. \qquad (54)$$

With the strategy (54) we must now complute the mean excess failures. If $p_A > p_B$ then after N trials which result in a_N, b_N failures

$$C_N = E(a_N + b_N = Nq_A) \qquad (55)$$

Referring back to (46) and (48), there are two regions to focus on:

$$\begin{array}{ll} 1. & C_\infty(p_A, p_B, e, f) = \lim\limits_{N\to\infty} C_N < \infty \\ 2. & C_\infty(p_A, p_B, e, f) = \lim\limits_{N\to\infty} C_N/N > 0 \end{array} \qquad (56)$$

In case 2, we also have

$$c_\infty = (p_A - p_B)\beta$$

$$\text{where } \beta = \lim_{N\to\infty} \frac{E(N_B = B_N + b_N)}{N} \tag{57}$$

is the probability that the sampling process will be "trapped" in the poorer drug B.

9. No Bias: Prototype

We now have to analyze the development of the process represented by (54). All of the information is contained in $P_N(A,B,a,b)$, the probability that after N steps, we will have had A,B successes a,b failures of the two drugs. Of course,

$$P_0(A,B,a,b) = \delta_{A,0}\ \delta_{B,0}\ \delta_{a,0}\ \delta_{b,0} \tag{58}$$

and since

$$A + B + a + b = N \tag{59}$$

the subscript N can be dropped. It will also be convenient to introduce

$$\begin{aligned} P^1_{ABab} &= \theta\left[(A+e+1)(b+f+1) - (B+e+1)(a+f+1)\right] P_{ABab} \\ P^2_{ABab} &= \bar{\theta}\left[(A+e+1)(b+f+1) - (B+e+1)(a+f+1)\right] P_{ABab} \end{aligned} \tag{60}$$

where $\bar{\theta}(x) = 1-\theta(x)$

$\theta(x)$ defined by (8). According to (54), P^1_{ABab} is the probability that A will be tried at the next step, P^2_{ABab} that B will be tried.

It then follows, just as in (31) but including the initial state, that

$$\begin{aligned} P_{ABab} &= p_A\, P^1_{A-1,Bab} + p_B\, P^2_{A,B-1,ab} \\ &\quad + q_A\, P^1_{AB,a-1,b} + q_B\, P^2_{ABa,b-1} + \delta_{ABab,0000} \end{aligned} \tag{61}$$

This is a four variable partial difference equation with nonconstant coefficients, a quite difficult object to deal with. However, assuming that it has been solved, the mean excess total failures is readily computed. Since trying B entails a mean excess failure of $p_A - p_B$ and the probability of trying B is P^2 of the previous step, we have at once

$$C_N = (p_A - p_B) \sum_{A+B+a+b<N} P^2_{ABab} \tag{62}$$

Let us start our examination of unterminated ethical strategies with the case of no initial bias towards success or failure:

$$\text{no bias:} \quad e = f = 0 \tag{63}$$

Since this is still a difficult analytical problem, we first consider the extreme

$$p_A = 1 \ (q_A = 0) \text{ with } p_B = p \ , \ q_B = q \tag{64}$$

as we did in (46). Since a remains 0, it can be dropped as an index, and (61) becomes

$$P_{ABb} = P^1_{A-1,Bb} + pP^2_{A,B-1,b} + qP^2_{AB,b-1} + \delta_{ABb,000} \tag{65}$$

with P^1 and P^2 given by (60) at $e = f = a = 0$. To include (60) automatically we introduce a generating function

$$P_{Bb}(z,t) = \sum_A P_{ABb} \, z^A \, t^{(A+1)(b+1)-(B+1)} \tag{66}$$

which converges in the complex plane for $|t| = 1$ at $|z| < 1$; and further decompose this via

$$P_{ABb} = P^1_{ABb} + P^2_{ABb} \text{ into } P_{Bb}(z,t) = P^1_{Bb}(z,t) + P^2_{Bb}(z,\tfrac{1}{t}) \tag{67}$$

According to (60) we thus have

$$\begin{aligned} &P^1(z,t) \text{ is a MacLaurin series in } t \\ &P^2(z,\tfrac{1}{t}) \text{ is a Laurent series in } t \\ &P^1(z,0) = P^2(z,0) \end{aligned} \tag{68}$$

On summing over A with the weight of (66), (65) reads

$$P_{Bb}(z,t) = zt^{b+1}\, P^1{}_{Bb}(z,t) + \frac{p}{t}\, P^2{}_{B-1,b}(z, \frac{1}{t})$$

$$+ qt\, P^2{}_{B,b-1}(zt, \frac{1}{t}) + \delta_{Bb,00} \tag{69}$$

If we are interested in C_∞ , it is convenient to sum (69) over all B. Setting $P_b(z,t) = \sum_B P_{Bb}(z,t)$ this yields

$$(1 - zt^{b+1})\, P^1_b(z,t) = qtP^2_{b-1}(zt,\frac{1}{t}) - (1-\frac{p}{t})\, P^2_b\,(z,\frac{1}{t})$$

$$+ \delta_{b,0} \tag{70}$$

It is not hard to show from (70) that $P^2_b(z,\frac{1}{t})$ is independent of z. In fact, a detailed proof is unnecessary: since a=0, drug B will only be tried if $A \leq B$ at every step of a chain of trials from B=0 on; hence, A=0. Taking z=0 and dropping the associated argument, we have

$$P^1_b(t) = qtP^2_{b-1}\ (\frac{1}{t}) - (1 - \frac{p}{t})\, P^2_b\, (\frac{1}{t}) + \delta_{b,0}$$

$$P^1_b(0) = P^2_b(0) \tag{71}$$

eq. (71) is easy to solve. We take the Laurent part

$$P_b^1(0) = qt\left(P_{b-1}^2(\tfrac{1}{t}) - P_{b-1}^2(0)\right) - (1 - \tfrac{p}{t})P_b^2(\tfrac{1}{t}) + \delta_{b,0} \tag{72}$$

use $P_b^1\ (0) = P_b^2\ (0)$ and sum over all b, dropping the associated index:

$$P^2(0) = qt\left(P^2\ (\tfrac{1}{t}) - P^2(0)\right) - (1 - \tfrac{p}{t})\ P^2(\tfrac{1}{t}) + 1$$

$$\text{where} \quad P^2(\tfrac{1}{t}) = \sum_{ABb} z^A\ t^{(A+1)(b+1)-(B+1)} P^2_{ABb} \tag{73}$$

Thus

$$P^2(\tfrac{1}{t}) = \frac{1-P^2(0)(1 + qt)}{(1 - t)(q- p/t)} \tag{74}$$

From (62)

$$C_\infty = qP^2(1) \tag{75}$$

which will be finite only if the numerator of (74) vanishes at t = 1:

$$P^2(0) = \frac{1}{1+q} \tag{76}$$

The remaining pole of (74) occurs outside of the region $|\frac{1}{t}| \leq 1$ only if $q > p$, and so we conclude on substituting (74), (76) into (75) that

$$p_B < \frac{1}{2} \qquad C_\infty = \frac{q_B^2}{(1+q_B)(1 - 2p_B)} \tag{77}$$

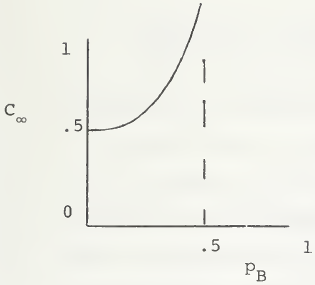

Figure 2

looking as shown. If $p_B > \frac{1}{2}$, only C_N/N has a limit, and an extension of the previous analysis shows that then:

$$p_B > \frac{1}{2} : \quad c_\infty = \lim_{N\to\infty} \frac{C_N}{N} = 1 - \frac{1-q_B + q_B^2}{1 - \frac{1}{2}q_B} \tag{78}$$

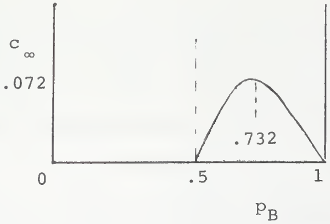

Figure 3

10. No Bias : General

We proceed to the general case of unbiased a-priori distribution: $f_0(p_A , p_B) = 1$. Our evaluation will still be made in terms of the asymptotic mean excess failures, with finite C_∞ signifying an asymptotically "safe" point in (p_A , p_B) space. From the viewpoint of the trapping probability β of (57), there is a region of the $(p_A p_B)$ space, including (1,0), in which $\beta = 0$ throughout. Crossing the boundary of this region one has a "phase transition" and β becomes non-vanishing, so that the sampling process can branch asymptotically to repeated sampling of either the better or worse of the two drugs. In this latter region, the mean excess failures depends upon the patient horizon N and the sampling parameters for optimal strategy should therefore depend on N as well. However, we will concentrate our attention in this lecture on the $\beta = 0$ regime, and hence on the total mean failures C_∞ .

Since our procedure, like any other, is most effective in the extreme case $p_A = 1$, $p_B = 0$, a sensible analysis should focus on the penalty exacted by non-vanishing values of $q_A = 1 - p_A$ and p_B . For this purpose, we may rewrite (61) as

$$P_{ABab} - P^1_{A-1,Bab} - P^2_{ABa,b-1} =$$

$$\delta_{ABab,0000} + q_A(P^1_{AB,a-1,b} - P^1_{A-1,Bab}) + p_B\,(P^2_{A,B-1,ab} - P^2_{ABa,b-1}) \tag{79}$$

and carry out an expansion with respect to q_A and p_B. This is not trivial, due to non-constant coefficients implicit in P^1 and P^2. Of course, part of the q_A , p_B expansion is already known, since (77) expands out to

$$C_\infty(0,p_B) = (1-p_B)[\tfrac{1}{2} + \tfrac{3}{4}p_B + \tfrac{11}{8}\,p_B^2 + \tfrac{43}{16}\,p_B^3 + \ldots] \tag{80}$$

Similarly, one can with considerably greater difficulty, solve the $p_B = 0$ case, obtaining

$$C_\infty(q_A,0) = (1-q_A)(1+q_A)\sum_{b=0}^{\infty} t_b^{*\,b+1}\,(q_A + t_b^*) \tag{81}$$

where $t^*_b = t_b(1)$ and $t_b(\lambda)$ satisfies

$$\begin{gathered} t - \lambda q_A = \lambda p_A t^{b+2} \\ t(\lambda) \to 0 \quad \text{as } \lambda \to 0 \end{gathered} \tag{82}$$

On expanding, we find

$$C_\infty(q_A,0) = (1 - q_A)\ [\tfrac{1}{2} + \tfrac{5}{4}q_A + \tfrac{11}{8}\,q_A^2 + \ldots] \qquad (83)$$

The cross terms, however, must be picked up by laborious appliecation of (79). Through second order, this results in

$$C_\infty(q_A,p_B) = (p_A - p_B)\ [\tfrac{1}{2} + \tfrac{3}{4}p_B + \tfrac{5}{4}q_A + \tfrac{11}{8}\,(p_B^2 + q_A^2) + \tfrac{21}{4}\,p_B q_A + \ldots\ ] \qquad (84)$$

How is an expansion such as (84) to be used? In particular, can we use it to estimate the boundary in $(q_A\ ,\ p_B)$ space at which the transition from untrapped to trapped sampling takes place? To do so, we convert (84) to a Pade type rational fraction expansion. First introduce an expansion parameter γ : $p_B \to p_B, q_A \to \gamma q_A$ and then convert the γ series to a continued fraction:

$$1 + a\gamma + b\gamma^2 + \ldots = 1 + \frac{a\gamma}{1 - \frac{b}{a}\gamma} + \ldots.$$

$$= (a + (a^2 - b)\gamma + \ldots)\ /\ (a - b\gamma + \ldots)\ .$$

In the present case, this yields

$$C_\infty(q_A,p_B) = \frac{1}{2}(p_A-p_B)\,\frac{\frac{3}{2}p_B + \frac{5}{2}q_A - \frac{1}{2}p_B^2 + \frac{7}{2}q_A^2 - 3p_Bq_A + \ldots}{\frac{3}{2}p_B + \frac{5}{2}q_A - \frac{11}{4}p_B^2 - \frac{11}{4}q_A^2 - \frac{21}{2}p_Bq_A + \ldots} \tag{85}$$

so that the transition boundary is given by

$$\frac{3}{2}p_B + \frac{5}{2}q_A - \frac{11}{4}p_B^2 - \frac{11}{4}q_A^2 - \frac{21}{2}p_Bq_A + \ldots = 0 \tag{86}$$

as shown. Despite the rough nature of the computation, this result compares very well with a numerical simulation that has been performed, using 100 runs of 1000 patients at each point of a (p_A,p_B) grid with grid spacing 0.1.

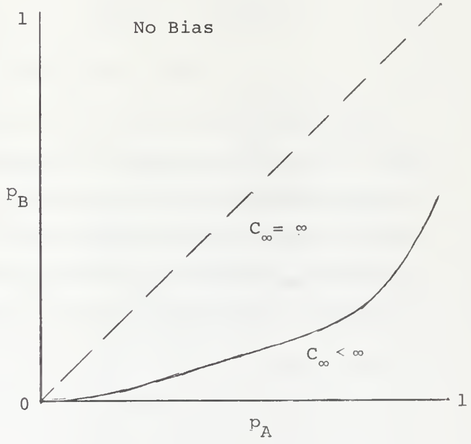

Figure 4.

At this point, let us pause to inquire as to how much has really been accomplished, compared for example to the unterminated PW scheme. Indeed, at $p_A = 1$, PW yield a finite C_∞, which is true for our fully ethical procedure at 0 bias only for $p_B < \frac{1}{2}$ (even then, PW has a lowever C_∞). But if $p_A < 1$, the story is very different. There, as we have seen in (48), only

$$\beta = \frac{C_\infty}{p_A - p_B} = \frac{q_A}{q_A + q_B} \tag{87}$$

is finite and $C_N \propto N$ (not due to trapping, to be sure, but rather to continued sampling of both drugs). A tabular comparison of PW and "zero bias ethical" shows that not only does the latter have an extensive region in which C_∞ remains finite, but when it is not finite, a smaller β value results through most of the (p_A, p_B) plane.

Table 3

100 β

N = 1000

No Bias

1

p_B

									34
								40	36
							42	39	29
						41	31	20	15
					32	19	9	3	3
				39	22	8	2	2	0
			32	15	7	4	1	1	0
		16	11	3	1	1	0	0	0
	9	4	1	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0

0 p_A 1

Table 4

100 β

N = ∞

P.W.

1

p_B

									0
								33	0
							40	25	0
						43	33	20	0
					44	38	29	17	0
				45	40	33	25	14	0
			46	42	36	30	22	13	0
		47	43	39	33	27	20	11	0
	47	44	40	36	31	25	18	10	0
47	44	41	38	33	29	23	17	9	0

0 p_A 1

11. Sampling with Bias

The problem with our no bias ethically motivated sampling procedure is quite evident from Figure 4. There is a non-vanishing probability β for the process to go asymptotically to repeated testing of the poorer drug unless the difference between p_A and p_B is substantial. This is true even at $p_A = 1$ (when $p_B > \frac{1}{2}$) , the one place when unterminated PW is totally effective. What happens is that an early fluctuation leading to a poor showing on trial of the better drug A establishes a reference that can be consistently exceeded by the poorer drug B. The remedy is clearly to assume initially a high level of performance by both drugs, so that an anomalously low reference will not be established. This is most readily accomplished in the present context by choosing a value $e > 0$ in the a-priori distribution f_0 of (51), resulting to be sure in a prolonged period during which the previous test results enter with a lower weight into each decision.

A quick estimate of the effect of $e > 0$ bias (with $f=0$) can be gotten by generalising the $p_A = 1$ case of (77) to $e > 0$. If this is done, one finds

$$p_A = 1 \quad : \quad C_\infty = \frac{q_B}{1 + p_B} \left\{ \frac{1}{\tau - 1} + \frac{1}{p_B \tau + 1} \right\} \tag{88}$$

where $\tau > 1$ satisfies

$$\tau^{e+2} - \frac{1}{p_B} \tau^{e+1} + \frac{q_B}{p_B} = 0 \tag{89}$$

Now C_∞ will remain finite until the large root of (89) falls to 1, yielding a double root of (89) at $\tau = 1$, which is readily seen to occur at

$$p_A = 1, \; p_{B\ \text{critical}} = \frac{e+1}{e+2} \tag{90}$$

In other words, the right intercept of the C_∞ transition curve is pushed up from $p_B = \frac{1}{2}$ at $e=0$ towards its maximum possible value of $p_B = 1$. But does the rest of the curve follow along? Suppose we choose $e = 10$, so that on a (p_A,p_B) grid of spacing 0.1, the whole right edge (below $p_A=p_B$ which we never include) will have finite C_∞, and again carry out numerical simulation with 100 sequences of patients at each grid point. For these finite samples, β is assessed (as it was before) by counting the ratio of sequences in which the last patient receives the drug B. With 200-patient sequences, continued testing by the poorer drug still occurs for $p_A - p_B = 0.2$, but closer to the asymptotic limit at $N = 1000$, β remains 0 until $p_A - p_B = 0.1$

Table 5

1

100 β
$N = 200$
Bias (10,0)

										5
									15	0
								15	1	0
							21	3	0	0
						21	2	0	0	0
					30	4	0	0	0	0
				34	5	0	0	0	0	0
p_B			15	1	0	0	0	0	0	0
		33	2	0	0	0	0	0	0	0
	21	1	0	0	0	0	0	0	0	0

0 p_A 1

Table 6

1

100 β
$N = 1000$
Bias (10,0)

										0
									3	0
								5	0	0
							5	0	0	0
						1	0	0	0	0
					0	0	0	0	0	0
				1	0	0	0	0	0	0
p_B			1	0	0	0	0	0	0	0
		0	0	0	0	0	0	0	0	0
	0	0	0	0	0	0	0	0	0	0

0 p_A 1

How do we measure the effectiveness of the strategy compiled in Tables 5 & 6? It is certainly not wonderful for $p_A - p_B \leq 0.1$. But of course clinical treatments are not highly stylized, nor are the patients equivalent, so that the success probability at each drug must be considered to have a distribution as well. For this reason, one finds in the literature that the smallest difference that one is called upon to discriminate is generally taken as 0.2. Hence a comparison at $p_A - p_B = 0.2$ is called for. This has been given for the optimal "traditional" strategies in Tables 1 & 2. In Table 7 we list the corresponding values of C_{200} and C_{1000} associated with the simulations of Tables 5 & 6.

Table 7

Mean Excess Failures for "Ethical" Sampling

Bias (10,0)

$p_A - p_B = 0.2$

N	$\bar{p} = 0.2$	$\bar{p} = 0.5$	$\bar{p} = 0.8$
200	6.7	4.5	1.6
1000	8.6	6.3	2.1

Consider first Table 1. Although the Biased Bayes (Ethical) criterion used in Tables 5 & 6 is unoptimized and for an infinite patient horizon, it is superior to "traditional" methods except for small population and low success probabilities . As we know, and observe from Table 1, $C_N \propto \ln_N$ for stopping rule methods so that the unterminated ethical procedure becomes better as N increases. On the other hand, if the expected values of p_A, p_B and horizon are poorly known, as in Table 2, the Biased Bayes is uniformly superior.

12. Conclusion

We conclude that a simple Bayes Criterion which satisfies reasonable standards of ethical practice, employed without a termination procedure, produces a success rate for large population which is at least comparable to that achieved by sampling methods not possessing the same ethical character. The question of optimal strategy for a given patient horizon, as well as modified treatment criteria for this purpose, is now under intensive investigation.

* Supported in part by the Department of Energy under Contract DE-AC02-76ER03077

References

[1] For earlier discussion of the general problem see:

Robbins, Herbert, "Some aspects of sequential design of experiments" Bulletin Amer. Math. Soc. 58 (1952) 527-535.

Zelen, M. "Play-The-Winner rule and controlled clinical trials" Journal Amer. Stat. Assoc. 64 (1969) 131-146.

[2] see e.g.

Sobel, M. & Weiss, G.H. "Play the winner sampling for selecting the better of two binomial populations" Biometrica 57 (1970) 357-365

[3] Upton, G.J.G. & Lee, R.D. "The importance of the patient horizon in sequential analysis of binomial clinical trials" Biometrica 63 (1976) 335-342

[4] Hoel, D.G., Sobel, M. & Weiss, G.H. "A two stage procedure for choosing the better of two binomial population" Biometrica 59 (1972) 517-523

[5] Berry, D.A. & Sobel, M. "Improved procedure for selecting the better of two Bernouilli populations." Jour. Amer. Stat. Assoc. 68 (1973) 979-984

[6] Fushimi, M. "An unproved version of Sobel-Weiss play-the-winner procedure for selecting the better of two binomial population" Biometrica 60 (1973) 517-523

[7] Simon, R. , Weiss, G.H. & Hoel, D.G. "Sequential analysis of binomial clinical trials" Biometrica 62 (1975) 195-200

[8] For earlier applications to deciding between several alternatives see:
Canner, P.L. "Selecting one of two treatments when the responses are dichotomous" Jour. Amer. Stat. Assoc. 65 (1970) 293-306

Feldman, D. "Contribution to the 'two-armed bandit' problem" Ann. Math Stat. 33 (1962) 847-856
Zaborskis "Sequential Bayesian plan for choosing the best method of medical treatment" Automation and Remote Control 37 (1976) 1750-1757